REMARKS

Claims 26-32, 34, and 36-40 are pending in the present application. Applicants are amending herewith Claim 27. Applicants are adding herewith new Claim 43. Applicants submit that support for these amendments can be found generally throughout the specification and that the amendments do not introduce new matter. Therefore, applicants submit that entry of these amendments is appropriate. Following entry of these amendments Claims 26-32, 34, 36-40 and 43 will be pending. Applicants respectfully request reconsideration of the present application in view of these foregoing amendments and the following remarks.

THE OFFICE ACTION

Claim 34 was provisionally rejected under the judicially created doctrine of obviousness-type double patenting in view of Claims 30, 32, 33, 35, 36, 38, 39 and 41 of copending application Serial No. 10/077,142, which is now U.S. Patent No. 6,908,910. Claims 26-32, 34 and 36-40 were provisionally rejected under the judicially created doctrine of obviousness-type double patenting in view of Claims 26-46 of copending application Serial No. 10/617,150, which is now U.S. Patent No. 6,930,128. Claims 27, 28 and 30 were provisionally rejected under the judicially created doctrine of obviousness-type double patenting in view of Claim 26 of copending application Serial No. 10/918,627. Claims 26-32, 34 and 36-40 were provisionally rejected under the judicially created doctrine of obviousness-type double patenting in view of Claims 1, 13 and 24 of copending application Serial No. 10/379,991, which is now abandoned. Claims 26, 27, 29-32, 34 and 36-40 were provisionally rejected under the judicially created doctrine of obviousness-type double patenting

in view of Claims 46-52 and 59-66 of copending application Serial No. 10/280,831. Claims 26-32, 34 and 36-40 were provisionally rejected under the judicially created doctrine of obviousness-type double patenting in view of Claims 26-46 of copending application Serial No. 10/789,471. Applicants are submitting herewith a terminal disclaimer as to some of the foregoing provisional rejections. Claims 27 and 28 were rejected under 35 U.S.C. §103(a) as being obvious and unpatentable over the article by Seegers et al. Applicants respectfully traverse this rejection.

THE REJECTION BASED ON OBVIOUSNESS-TYPE DOUBLE PATENTING

As discussed above, various claims of the present application were rejected under the judicially created doctrine of obviousness-type double patenting in view of application Serial Nos. 10/077,142; 10/617,150; 10/918,627; 10/379,991; 10/280,831; and 10/789,471. Application Serial No. 10/077,142 is now U.S. Patent No. 6,908,910. Application Serial No. 10/617,150 is now U.S. Patent No. 6,930,128. Application Serial No. 10/918,627 is pending with no claims presently allowed. Application Serial No. 10/379,991 is abandoned. In application Serial No. 10/280,831 the issue fee has been paid and assignment of a patent number is imminent. Application Serial No. 10/789,471 is pending with no claims presently allowed.

With respect to the foregoing provisional rejections based on Application Serial Nos. 10/918,627; 10/379,991; and 10/789,471, since these rejections are provisional rejections and since the claims in those application do not appear that they will mature into a patent prior to the claims of the present application, it is submitted that no response to these provisional rejections is required at this time. With respect to U.S. Patent Nos. 6,908,910 and 6,930,128; and application Serial No. 10/280,831, applicants are submitting herewith a terminal disclaimer in

compliance with 37 CFR 1.321. Applicants are also submitting herewith the fee required by 37 CFR 1.20(d). Applicants submit that the filing of the terminal disclaimer overcomes the present provisional rejections. Accordingly, applicants respectfully request that the provisional rejections based on U.S. Patent Nos. 6,908,910 and 6,930,128; and application Serial No. 10/280,831 be withdrawn.

THE REJECTION BASED ON 35 U.S.C. §103

The Claims 27 and 28 were rejected under 35 U.S.C. §103(a) as being obvious and unpatentable over the article by Seegers et al. The article by Seegers et al. discloses the antimitotic effect of 2-methoxyestradiol on MCF-7 and HeLa cell lines **in vitro**. The rejection states that the instant claims differ from Seegers et al. in that the present claims are directed to a method of treating humans or animals. The rejection further states that it is routine in the art to extrapolate the effect of a compound from in vitro studies to in vivo use. The examiner concludes that a skilled artisan would have a reasonable expectation that 2-methoxyestradiol would be antimitotic when administered in vivo, and, thus, would be useful in treatment of solid tumors, such as breast and cervical cancers. The rejection states that the motivation for such use would be based on the teaching of the antimitotic effect of 2-methoxyestradiol in MCF-7 and HeLa cells as taught by Seegers et al. Applicants respectfully disagree.

Applicants are amending Claim 27 herewith to specify that the claim is a method of treating humans, rather than humans and animals. Applicants disagree that there is sufficient information in Seegers et al. to give a skilled person a reasonable expectation that 2-methoxyestradiol would be useful in treating humans for cancers, such as breast and cervical cancers. While in vivo testing is certainly sufficient to satisfy the disclosure and enablement requirements of 35 U.S.C. §112, applicants submit that a different test is applied for a rejection

based on 35 U.S.C. §103, that test being the reasonable expectation test. Applicants submit that the only reasonable expectation that a person skilled in the art would have based on Seegers et al. is that 2-methoxyestradiol is a cytotoxic poison for MCF-7 and HeLa cells in vitro. Applicants submit that Seegers et al. might suggest to a persons skilled in the art that it would be obvious to try the use of 2-methoxyestradiol in the treatment of animals and perhaps humans with breast or cervical cancer. However, "obvious to try" is not the appropriate test under 35 U.S.C. §103.

As stated above, a positive in vitro test does not give a person skilled in the art a reasonable expectation that the compound would be useful in treating humans. Humans are very complex creatures. The normal testing hierarchy is to go from in vitro testing to testing in animal models, such as mice or rats. If the animal models are positive then the tests may proceed to testing in humans. However, it is well known that many promising cancer drugs which showed positive test results in mice or rats did not prove to be effective in humans. Therefore, a person skilled in the art in the field of cancer treatments would knows that animal models are not sufficiently predictive of results in humans to give them a reasonable expectation of success in treating humans for cancer. In vitro tests are one step further away from humans than animals, which makes in vitro testing even less predictable of results in humans.

If finding a cure for cancer was as simple as finding a compound that tested positively under in vitro testing, a cure for cancer would have been found years ago. However, in the field of cancer treatments, it is believed that in vitro testing is not known to be predictive of results in humans. The examiner states that in the medical art it is routine "to extrapolate the effect of a compound from in vitro studies to in vivo use." However, the examiner offers no evidence that a person skilled in the art would recognize that an in vivo antimitotic effect on

MFC-7 or HeLa cells is in any way correlated to the treatment of solid tumors in humans. Applicants submit that the burden is on the examiner to produce such evidence.

The attached article Lu et al., "From Correlation to Causation to Control: Utilizing Preclinical Disease Models to Improve Cancer Target Discovery," PRECLINICA (March/April 2003) suggests that some in vitro tests have low predictability in animals and even lower predictability in humans. This article states:

> One clear example, reviewed recently (30), can be found in the hereditary breast cancer gene *Brca1* and *Brca2*, which may be excellent for diagnosis, but so far have failed for therapeutics. Unfortunately, the problem is exacerbated by low predictability of in vitro cell models and model organisms for gene and protein roles even in animal disease models and even lower predictability for human disease. Validation of candidate drug target genes and proteins is becoming focused on whether or not they have the ability to actually control an ongoing disease process and reverse the pathology or its symptoms. This is largely due to a number of complex interactions of multiple cell types that result in disease pathology. Moreover, many proteins lie in pathways that can be detected or validated in active disease tissue.

Lu et al. at page 36. Thus, the foregoing demonstrates that the examiner's sweeping generalization that it is routine in the art to extrapolate the effect of a compound from in vitro studies to in vivo use is simply not true. This is particularly true for in vitro tests that have no recognized correlation to human disease, such as those in Seegers et al. Again, applicants request that the examiner provide evidence that the prior art recognized that the in vitro antimitotic effect in Seegers et al. had a correlation to the treatment of solid tumors in humans.

The attached article, J.E. Tomaszewski, "The Importance of Being Able to Predict Human Sensitivity," NIH Summit Workshop on Predictive Drug Toxicology (June 15-17 2004), shows that of 5,000-10,000 drug candidates, only 2,500 enter preclinical testing (*i.e.*, animal testing) and only five enter clinical testing (*i.e.*, testing on humans) and only one is approved by

the FDA. That means that only 0.05%-0.1% of drug candidates, which are typically identified by in vitro testing, ever make it to testing in humans and 0.01%-0.02% are deemed effective in humans by the FDA. Thus, it can hardly be said that it is routine in the art to extrapolate the effect of a compound from in vitro studies to in vivo use. The foregoing suggests that such is not the case in drug discovery.

The examiner also states that Seegers et al. would lead the skilled artisan to the reasonable expectation that 2-metoxyestadiol would be antimitotic when administered in vivo, and, thus, would be useful in the treatment of solid tumors, such as breast and cervical cancers. However, there is no prior art teaching of record that an antimitotic effect in humans is useful in treating solid tumors. It is submitted that the examiner is using hindsight reconstruction to connect the in vitro antimitotic effect of Seegers et al. to the treatment of solid tumors in humans. Such an approach to the issue of obviousness, of course, is improper.

Furthermore, it cannot be determined from the results shown in Seegers et al. whether the cytotoxic effects of 2-metoxyestadiol would be selective for dividing cancer cells or would be cytotoxic to all human cells; in which case it would be a poison, not a cancer treatment. Without some degree of selectivity, a cytotoxic agent would not be useful for treating solid tumors in humans. However, selectivity in humans cannot be determined from the in vitro results of Seegers et al. Again, it is submitted that the examiner is using hindsight reconstruction to conclude that 2-metoxyestadiol would have a selective effect in humans since there is no teaching in the prior art of this effect.

The disclosure of Seegers et al. is simply insufficient to give a skilled artisan a reasonable expectation that 2-metoxyestadiol would be effective in treating solid tumors in humans. Therefore, it is believed that the examiner's conclusions regarding the alleged

obviousness of Claims 27 and 28 are incorrect and that the rejection of those claims under 35 U.S.C. §103 is inappropriate. In view of the foregoing, applicants submit that the rejection of Claims 27 and 28 under 35 U.S.C. §103 should be withdrawn.

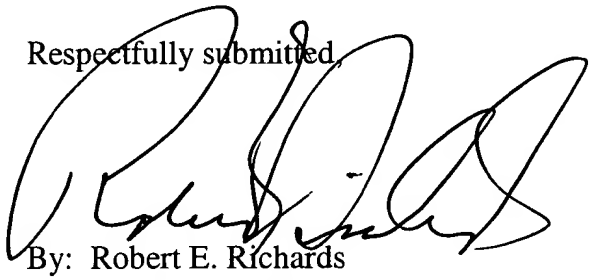
THE NEW CLAIM

Applicants are adding herewith a new Claim 43. This claim is directed to a method of inhibiting excessive or abnormal stimulation of endothelial cells in a human or an animal. Applicants submit that support for this new claim can be found throughout the specification in general, and particularly at page 3, line 8 of the present specification.

CONCLUSION

In view of the foregoing remarks and amendments, applicants submit that the claims define patentable subject matter over the prior art and are in condition for allowance. A Notice of Allowance is therefore requested and such action is respectfully solicited. If the Examiner believes that any informality remains in the application that may be corrected by Examiner's Amendment, or there are any other issues which can be resolved by telephone interview, a telephone call to the undersigned attorney at (404) 745-2408 is requested.

Respectfully submitted,



By: Robert E. Richards
Reg. No. 29,105

Kilpatrick Stockton LLP
Suite 2800
1100 Peachtree Street
Atlanta, GA 30309-4530
Tel. (404) 815-6500
Our File: 05213-0493 (43170-253692)

From Correlation to Causation to Control: Utilizing Preclinical Disease Models to Improve Cancer Target Discovery

Patrick Y. Lu, Frank Y. Xie, Puthupparampil Scaria, and Martin C. Woodle
Intradigm Corporation, Rockville, MD, USA

ABSTRACT

Adaptation of genomic and proteomic methods to preclinical animal cancer models promises to revolutionize cancer drug target discovery. Preclinical cancer models uniquely permit focusing on the salient aspect of a drug target, genes, and proteins associated with changes in established tumors. We have identified a direct and rapid manner to use animal cancer models to identify proteins associated with efficacy. The method commences with altered gene expression of tumor-controlling proteins in specific tumor pathways to perturb the growth of established tumors. The resulting tumor perturbation is used to obtain a differential measure of the genes and proteins associated with tumor growth dynamics. Recent advances in efficient nucleic acid delivery tools with low background biological activity enable the concept. First, methods for systemic expression of plasmids at the liver permit separation of gene expression activity and tumor response. Local administration of genes and oligonucleotide inhibitors without proteinaceous viral vectors permit local effects with minimal background biological activity. The second step of the concept is application of genomic and proteomic technologies to tumors with growth rate changes to reveal the genes and proteins correlating with efficacy processes. The third and final step utilizes nucleic acid delivery to alter the expression of efficacy-associated genes in the same or different animal cancer models to validate the genes for efficacy. The targets validated in this manner can then be studied with the standard genomic, proteomic, and pharmacological methods for toxicity in a narrow focused manner to complete their validation as cancer drug targets.

INTRODUCTION

Sequencing the entire human and mouse genomes has added massive new challenges to struggling functional genomics efforts, but interest remains strong in genomics-based drug discovery. For example, the pure sequence information has mushroomed, overloading computational processing methods. Hypotheses abound on how to utilize all the information and techniques of genomics and proteomics. A consensus does exist that the first step in utilizing the genomic sequence information for therapeutic development is the identification of new or better drug targets. Drug targets are the key genes and proteins that more or less single-handedly control the underlying pathology of a disease (i.e., the "gatekeepers" of pathological processes) (1,2). The result of this consensus is considerable focus onto drug target discovery with the objective of developing the most efficient methods of finding the gatekeeper pathways and pathway members.

The problem is very much like trying to find needles in a haystack; both will poke the skin, but only the needle is strong enough for sewing. It is not yet clear how many of the approximately 40,000 human genes will prove to be strong drug targets. The current set of approved drugs utilizes only about 500 proteins, implying that, since there are many human genes, many targets are waiting to be discovered. The problem is even greater as a result of alternative mRNA splicing and posttranslational modifications such as phosphorylation. Consequently, the complete list of all proteins and, thus, drug target candidates, is more likely 10 times as many as the number of genes, up to 500,000 or perhaps even higher. Sorting, prioritizing, and validating the best drug targets out all of these possibilities is a daunting task, and moreover, it must be repeated for each indication.

A common approach to solve the problem of finding all the drug targets out of all the human genes is to develop large-scale high-throughput studies with the ultimate goal of genome-wide searching. The obvious consequence of increasing the input scale of a search is production of enormous amounts of data. Unfortunately, the data produced are difficult to sort. Therefore, to easily reveal the few powerful gatekeepers requires massive computation capabilities (3–8). The field is struggling to identify the best methods for the filtering process. Consequently,

identifying the best search criteria (i.e., determining the best question to ask when looking through the enormous data sets) is one of the major focuses of ongoing efforts to solve the problem.

Simplified biological systems amenable to high-throughput assays are used in conventional target discovery, given the enormous pool of candidate drug targets and reliance upon high-throughput sorting. In many instances, preliminary candidate targets identified from these simplified systems are used for drug discovery, even though they are not fully validated for an ability to control disease pathology (9), which requires animal disease models and eventually human clinical trials. In fact, in some cases, drug discovery has been used to help discover and validate targets for involvement in disease processes (10,11). Considerable effort to improve this approach is being directed to construct high-throughput assays that will generate an understanding of the interactions between the many proteins that make up the pathways and, in the largest sense, build system-wide assemblies of protein interactions called "system biology". Most recently, this has culminated in efforts engaged in designing computational sorting of these large data sets, so as to identify the processes and of course the players in those processes, the "computational systems biology" (12). Many approaches have the goal of finding computational systems and methods to apply to the large data sets so as to identify the relationships between players (i.e., protein targets) (13). These and many other efforts are ongoing to determine all protein–protein interactions, all signal transduction pathways, all transcription factors and their regulation, etc., and the result is an exponential increase in the data sets and their size. Unfortunately, despite the ambition and thoroughness of the large-scale envisioned (building the macroscopic system by gathering all information on all of the microscopic elements and all of their interactions), the value of these approaches is still not clear. They have yet to yield information needed to tweeze out the key controlling elements that switch pathology off (i.e., the highly prized drug targets).

A more important problem, beyond the massive size of the data, is making sure the data sets actually contain the desired information. A growing effort is being applied to the design of the biological assay system to improve content. The concept is that the source of the data is equally, if not more, important than the sorting algorithm. It appears that lack of attention to this key aspect has been the most important factor limiting successful application of genomic and proteomic methods. Approaches to address this problem include generating data sets containing gene and protein information relating to stage of disease (14) and, more recently, relating to the cause of disease (15,16).

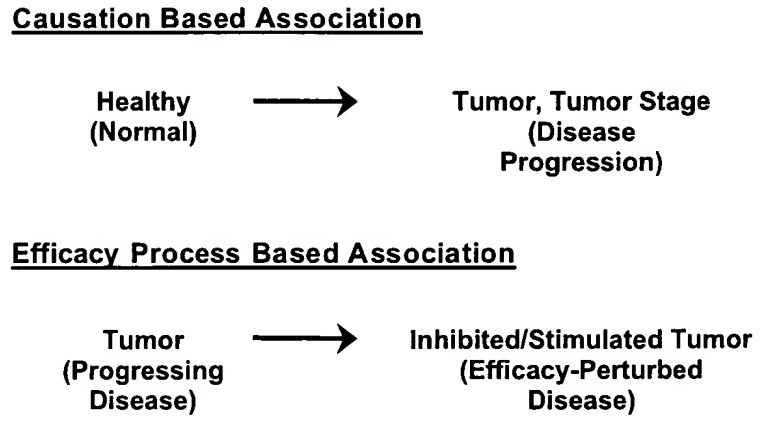


Figure 1. Comparison of disease correlation and disease causation biological systems with efficacy process biological systems for target discovery. The correlation or causation systems start with healthy or normal tissues for the differential, while the efficacy process system starts with control disease, that is unperturbed progressing disease. In the case for cancer targets, the efficacy process is the growth rate of tumors, reduced metastasis, etc.

Our concept is that preclinical animal models of disease, preferably clinically relevant disease models, can be used to improve data set content. Specifically, preclinical disease models can be used to generate data sets enriched with disease dynamics information. It is the dynamics that is needed for therapeutic effects and the gene and proteins associated with the dynamics involved in control of the disease pathology that are needed as drug targets. Here, we describe a concept for development of animal models for the generation of gene and protein data sets enriched for answers to the question; "what genes and proteins are in pathways that control the pathology". In other words, in order for therapeutic intervention to be successful, it must control the pathological process or at least the symptoms that result. To fulfill that search criteria, the data set needs to contain, if not enriched in, data that relates to controlling pathological processes: the inhibition of undesirable processes and the enhancement of desirable processes. We call these desirable and undesirable disease-controlling processes "efficacious processes". If our hypothesis that preclinical models showing efficacious processes can improve the quality of the data set, and potentially the concentration of information for use by the sorting processes, then this approach should

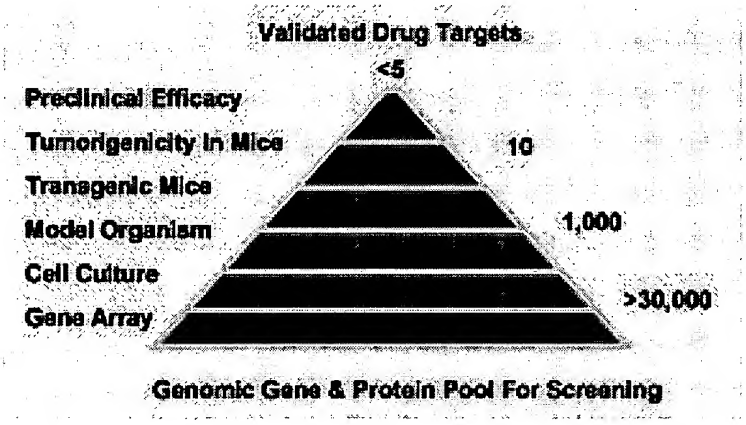


Figure 2. Reduction in the number of candidate genes and proteins considered as cancer drug targets as a result of standard genomic and proteomic sorting processes relying on preclinical animal models only at the last steps.

prove to provide data sets that are enriched with the desired information.

This report describes the use of preclinical animal disease models as a system that can be generated to test the hypothesis. The data set must be generated from systems capable of eliciting efficacious processes, either undesirable processes or desirable processes, or preferably both. This means systems are needed that not only contain the pathological state, but where the pathological state can respond and produce efficacious processes (i.e., systems with active disease pathology). Generally, active disease pathology for manipulation is only found in preclinical disease models. Therefore, we propose the use of preclinical animal disease models, since they clearly enable production of data sets containing efficacious process information. To test this concept, we have designed systems in which animal cancer models can be perturbed using nucleic acid delivery so as to elicit specific changes in protein expression that induce the dynamic processes desired, namely inhibited and stimulated tumor growth. A comparison of this disease dynamics-based concept with the conventional approaches based on correlation with disease or more recent efforts using causation is shown in Figure 1. It is our goal to design systems enriched for discovery of efficacious pathways, and with the biological systems available today, this requires preclinical disease models. An additional potential benefit, by identification of efficacious pathways and the genes and proteins that compose them directly in preclinical models of disease, is that the total number of animals required for development of efficacious drugs should decline as a result of a reduction in preclinical studies on poor targets and the drugs developed for them. We first review the nature of disease correlation- and causation-based biological systems and then describe the design of such a system for cancer to test this hypothesis.

BIOLOGY: DISEASE CORRELATION

The initial and most prevalent approach to forming biological systems for the generation of biological data sets relies on correlation with disease. Many biological systems have been devised for identification of genes and proteins correlating with disease or to segregate them between disease and healthy states. The most common example is the generation of data sets from cells and tissues classified as either healthy or diseased (6,17). The critical issue to data generation from these biological systems is their dependence on rapidly advancing the throughput of genomic and proteomic tools in order to obtain more and better data relating the segregation of genes and proteins between these two types of samples.

The most important genomic tool yet developed is the measurement of expressed RNA profiles using microarray determination of mRNA sequences (6,18). This technology has been developed, successfully enabling high-throughput screening of the active gene expression occurring in samples of cells and tissues. The resulting data sets contain measures of the level of each individual mRNA sequence in the mRNA pool prepared from each cell or tissue sample tested. The "bandwidth" of the genes in the data set is determined by the array size and content. While human genome-wide microarrays generally are not available commercially, some have been developed for over 30,000 mRNA sequences. Now microarray technology has become a well-established and reproducible robust method to obtain gene expression information from a large number of biological sample types, including human tissues. The primary application of microarray expression profile measurement has been to assemble data sets that allow identification of genes that correlate with certain biological traits or conditions. Furthermore, this use has resulted in large databases of human pathological and healthy tissue gene expression, some of which are commercially available (6). These databases are providing increasing amounts of data on the correlation of gene expression with specific pathological tissues, and they can be searched (i.e., sorted) for individual genes whose expression is

correlated with that condition. Consequently, large-scale searching of these databases, assisted by new and better methods of sorting the data, is being used to identify gene expression patterns that correlate with disease. One recent example is the determination of the correlation of known angiogenic gene expression with a few specific cancer types (19).

The successful establishment of large genomics databases has inspired efforts to expand into proteomics databases. This has led to an extensive effort to achieve similar databases containing protein information, including posttranslational modification of proteins such as phosphorylation state (20,21). Initially, 2-dimensional gel separation methods were developed to separate and identify the specific proteins and their phosphorylation state in a cell or tissue extract. Like microarrays for gene expression, 2-dimensional gels have been successfully developed into a reproducible and robust method. More recently, alternative separation methods with more specific activities have been developed (21). Unlike measurement of gene expression levels on a microarray, however, determination of each protein after separation is much more difficult and complicated. Currently, mass spectroscopy is one of the most commonly applied methods to identify each protein and reveal important information such as phosphorylation (21). When these methods for protein measurements are applied on large numbers of samples of cells and tissues, including human tissues, in the same way that microarrays have been applied for gene expression, they will greatly expand the already large databases of information that provide correlation of proteins with healthy or disease states. As these databases are combined, great levels of information on the correlation of specific gene expression and proteins with disease will be available for bioinformatic sorting. This process has already begun (e.g., 19,20).

IMPROVED BIOLOGY: FROM CORRELATION TO CAUSATION

A recent effort to improve the biological systems and resulting data sets is directed toward extending from a simple disease correlation to correlation with stage of disease. This helps to capture information on genes and proteins associated with disease progression, which is essentially information on disease causation. Much of this expansion in biological data generation utilizes disease tissue samples with pathological characterization as to

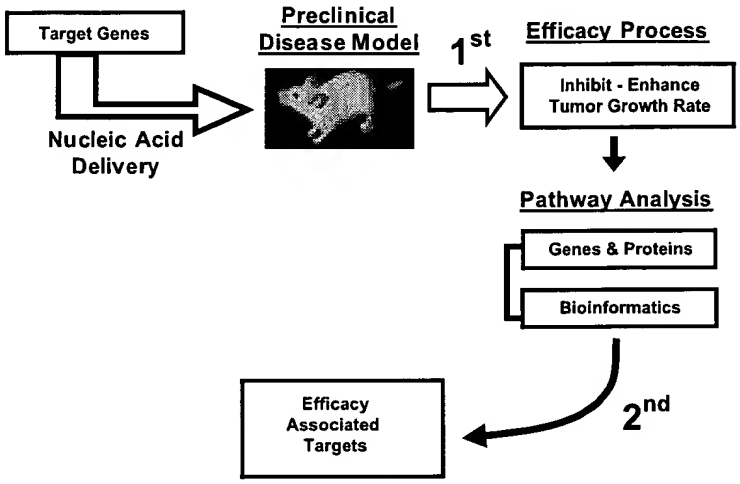


Figure 3. Use of nucleic acid delivery into preclinical animal models of cancer to induce efficacy as a first step in cancer drug target identification. The nucleic acid delivery achieves a gene-specific inhibition or expression to induce macroscopic pharmacological processes associated with efficacy stimulated and inhibited tumor growth rate. Pathway analysis of genes and proteins in the nucleic acid-perturbed tumors is used to identify genes and proteins whose expression or activity correlates with the induced alteration of tumor growth rate. The differential between control tumors and those with altered growth rate gives candidate targets for therapeutic intervention.

the stage of disease and includes a growing use of human tissue samples (e.g., 6). However, important enhancements have been made by two technological advances: (*i*) laser capture microdissection (LCM), which reduces the tissue and cells collected for expression profiling to very specific regions of a tissue section that are ascribed to a specific disease state (22–25); and (*ii*) increased sensitivities of assays that reduce sample requirements for gene and protein profiling. In the case of gene expression profiling, the increased sensitivity comes through the introduction of amplification techniques, but also may introduce errors since some signals may be preferentially amplified (26,27). These improvements increase the quality of the data set, in that the information provides better selectivity of genes and proteins associated with disease or with stage of disease. One recent approach has adapted a preclinical xenograft cancer model to derive tissues with increasing tumorigenicity for gene profile measurements, providing a data set of gene expression changes as a function of disease progression (14). Thus, there is a continuing effort striving to develop biological systems that can provide improved data sets with an emphasis on developing systems, enabling better biological data correlating genes and proteins with stage of disease.

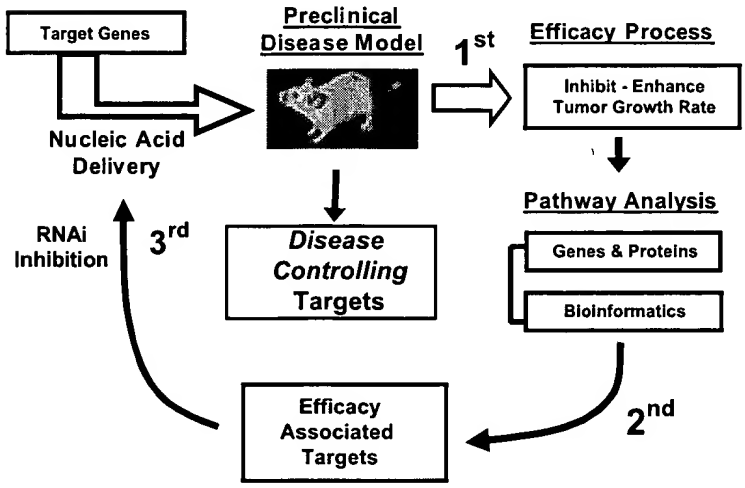


Figure 4. Use of nucleic acid delivery into preclinical animal models of cancer to sort efficacy-associated target genes using RNAi inhibitory oligonucleotides as a validation for target control of tumor growth. The nucleic acid delivery achieves a gene-specific inhibition within the tumor. When the reduced expression of that single gene induces the macroscopic pharmacological processes that alter tumor growth rate, then that gene and the protein it expresses is identified as a disease-controlling target.

Another growing effort is directed to developing biological systems that provide data sets derived from biological systems providing direct measures of disease causation. These efforts essentially represent a return to conventional biology but employing modern tools and methods. The goal is to achieve a further improvement in the generation of data sets, in terms of the biological information they contain, by developing systems that enable the generation of data sets derived not on correlation or even stage of disease, but specifically on causation of disease (e.g., 15,16,28,29). They also promise enrichment for specific genetic defects responsible for genetic disorders. The effort to obtain disease causation information is motivated by a realization that many genes and proteins may correlate with a disease condition or stage of disease but will lack a active role in the disease process, pathology, or generation of symptoms and, thus, have no value for drug intervention in the disease. It is also based on the original approach to biology, starting with phenomenological observations and attempting to identify the underlying mechanism (i.e., genetics) and thus has been called "forward genomics" (e.g., 28). Note that effort to reveal proteins correlating with disease and stage of disease still has tremendous value, as these proteins are excellent candidates for diagnostics, but this selection criteria is not as useful for drug targets to reverse disease pathology. Thus, in an effort to address this limitation of disease correlation, approaches have been developed to increase the value of the biological data, namely by identifying the proteins that play a role in causing disease and pathology.

A very successful approach to identify oncogenes and other disease-causing genes (those genes whose mutation or overexpression causes or contributes a step in the development of cancer or other diseases) has been by exposure of animals to mutagenic agents and then identifying the mutated genes in animals that develop disease (e.g.,15,16). The challenge is identifying which specific genes are mutated and, most importantly, insuring that only the identified mutations are actually responsible for the development of the disease. One solution to this problem utilizes labeled virus that integrates into the mouse genome and as a result either inactivates or overexpresses genes where the integration occurs (16). The label in the virus allows rapid determination of the locations where the virus has integrated, helping to address the first problem. In cancer, this method appears to have an inherent ability to reduce or eliminate the second problem. Cells with mutations that do not contribute to the formation of cancer appear to be eliminated, so as to deselect all virus-induced mutations without oncogenic effects. The resulting tumors appear to contain largely or potentially only labeled virus-induced oncogenic mutations. Similar approaches have been developed using chemical mutagens that cause single point mutations in DNA and searching algorithms to identify the causative mutations (15,29). These approaches, through development and use of improved virology, molecular biology, bioinformatics, etc., have revolutionized the speed of the original "forward genomic" biology methods to identify the mechanism of biological phenomena (i.e., the cause of disease biology). The constructed biological systems and resulting data sets, being enriched with information on causation of disease, offer a significant advance beyond data sets based on correlation with disease or disease stage.

DISEASE BIOLOGY: THE CHALLENGE OF GETTING FROM CAUSATION TO DISEASE CONTROL

For discovery of therapeutics to control disease, a need to go beyond identification of disease-causing genes and proteins has begun to be recognized (30). The biological incentive and desired outcome of drug intervention is just that, the intervention in the

disease process or pathology. Thus, improvement of data sets to include information on causation of disease, unfortunately, still does not address the key to identify and select drug targets, namely biological control functions. The evidence is growing that the identification of disease-causing genes has not proven to facilitate identification of viable small molecule drug targets. One clear example, reviewed recently (30), can be found in the hereditary breast cancer genes *Brca1* and *Brca2*, which may be excellent for diagnostics, but so far have failed for therapeutics. Unfortunately, the problem is exacerbated by low predictability of in vitro cell models and model organisms for gene and protein roles even in animal disease models and even lower predictability for human disease. Validation of candidate drug target genes and proteins is becoming focused on whether or not they have the ability to actually control an ongoing disease process and reverse the pathology or its symptoms. This is largely due to a large number of complex interactions of multiple cell types that result in disease pathology. Moreover, many proteins lie in pathways that can only be detected or validated in active disease tissues.

One approach to address drug target discovery and validation in animals has been the use of knock-out transgenic mice, in which a single gene is blocked from expressing in all cells of the animal (30,31). As these mice lack activity of a single specific gene, they provide evidence predicting the biological activity and consequences of a drug that inhibits the protein produced by that gene. The development of knock-out technology has been heralded as a method that can greatly facilitate high-throughput target selection and validation (31), in part through establishment of large animal farms based on libraries containing up to 5000 different knock-outs (30). Nonetheless, a couple of problems have yet to be solved in the use of transgenic mice to generate genome-wide biological systems for the production of large data sets. First, the entire set of animals that can be produced, as well as low-throughput at high cost, is far less than adequate to address all genes and proteins. This primarily results from many genes being required for early development. For adulthood however, they may yet play key roles in pathology and thus represent drug targets. Since these candidate targets are essential at one time or another, their knock-out is problematic or lethal. Nonetheless, the throughput of constructing knock-out mice has been increased to enable their generation for many thousands of genes (30), but the rate and cost still limits the application to all possible genes and certainly all protein forms from those genes. A more important problem is that the loss of a single gene in every cell from the fertilized egg to the adult simply predispositions the knock-out mice to either have an increased or decreased susceptibility to disease or to the progression of disease. The technique cannot provide a direct emulation of the effect a drug will have by inhibiting the target protein after the disease pathology is established and whether this inhibition will reverse the disease process. Such prediction for drug activity requires the reduction of the candidate gene and protein after the disease is established, thereby emulating the control or reversal of the disease pathology when a drug inhibits the target protein. Note that, classically, drugs inhibit their target, but the biotechnology-based pharmaceuticals expand this to include the addition of protein activity as a potential mechanism of therapy (as in the example of erythropoietin). One solution to this need to knock out the gene after adulthood or after the establishment of disease may come through advancements in mice knock-out technology, such as "conditional knock-outs", where the gene can be turned off conditionally (32). This advancement looks promising but is still very experimental and may be difficult to achieve the required high-throughput.

Continuing advancements are being made in animal disease models. Efforts are being made to develop simpler high-throughput biological test systems that recreate the properties of a human tissue that can propagate disease pathology. However today, the best biological system providing information on genes and proteins associated with disease control (i.e., reversal) are established, namely the preclinical disease models. The current repertoire of preclinical disease models is limiting in several manners, hence the extensive efforts to establish alternative biological test systems. While many preclinical models have demonstrated clinical relevance, many clinical indications lack good relevant animal models. Another problem is the relatively high cost and low-throughput of working with animals versus cell culture. Nonetheless, the biological advantages make the use of preclinical disease models an essential step in drug development, as shown in Figure 2, and the issue is largely at what stage to use them. Thus far, the use for gene and protein screening has been very limited, in part due to the low-throughput and in part due to the technical challenges of working with such large complex biological systems.

One challenge of using active preclinical disease models for gene and protein screening is getting clear information on specific gene functions (i.e., effects of their inhibition or enhancement) in the active pathological tissues. Recent progress has been described by collecting tumor tissues from animal models at different stages of the tumor proliferation and progression for genomic analysis (14), but again that provides largely oncogenesis information, albeit important and difficult to obtain information, such as the genes involved in the conversion of tumor cells to a metastatic phenotype.

The key problem of using animal disease models has been the lack of adequate methods to selectively inhibit or enhance specific genes in the disease tissue. One approach to this has been through the use of antisense inhibitors, but the achievement of antisense activity in animals has been limited (33,34). Further, acceptance of antisense for animal disease model application appears to await results from ongoing clinical development. However, recent rapid advances in gene-specific inhibition have been developed using double-stranded RNA (dsRNA), what is often called RNA interference (RNAi), including viral vector expression of dsRNA in animals (35). RNAi may improve animal disease model genomics but has only recently been described in animals. On the other hand, numerous applications of vectors to express genes and enhance their levels in specific tissues have been developed to support clinical gene therapeutics, with an effort continuing in cancer gene therapy heavily relying on selective viral replication (36,37). One approach to use gene expression for such gene function analysis has been proposed using very efficient liver delivery, resulting in strong expression of proteins secreted into the blood that can then act at implanted tumors (38). A key element of the challenge is that the amplitude of the inhibition or overexpression of the gene clearly must be large in order for a significant effect on phenotype to be observed. Nonetheless, this requirement appears to be a bigger challenge for clinical therapeutic responses than for the identification of gene function in a disease pathway. This may be particularly true for several classes of genes and proteins, such as secreted growth factors, upstream members of pathways, and in particular, for gatekeepers that are amplified by secondary factors in the pathways. Interestingly, these are the very same factors that promise to be the most efficacious drug targets.

Extension of high-throughput methods to animal disease models using libraries of gene constructs and simple administration may be possible with current technology, but appears to be a Herculean task. Alternatively, even using a relatively low-throughput expression or inhibition of specific target genes in combination with traditional genomic and proteomic profiling technologies may allow generation of biological data sets highly enriched with the pathological processes required for disease control and reversal. We have developed, and describe here, a concept for combining these two technologies that we call efficacy-linked genomics.

EFFICACY PATHWAY-LINKED GENOMICS: SCREENING DISEASE CONTROLLING PROCESSES

The factors associated with pathological pathways that control disease dynamics should be more easily identified in active preclinical disease models that are undergoing those dynamics. This is quite distinct from the well-accepted use of pathway analysis on isolated tissue sections as illustrated in Figure 1. To achieve a biological system that can address this difference, we have developed a concept using established (i.e., active) animal cancer models to obtain the essential dynamic information. The concept is to trigger efficacy processes in preclinical cancer models and then use the perturbed tissues for gene and protein analysis. The steps of the process are summarized in Figure 3. By using nucleic acid delivery into the preclinical animal model, a selected gene can be selectively up- or down-regulated so as to trigger efficacy processes and result in macroscopic changes in the established disease tissue. For this method to be effective, the perturbation needs to be selected either so that it triggers an observable effect on the disease or, alternatively, the approach can be used to obtain an understanding of pathways and functions of unknown genes that are perturbed. Another interesting use is with combination of drug treatments and perturbation of the gene. A second key requirement is that the nucleic acid delivery method needs to be capable of inducing a significant effect on the gene expression level in a manner that will trigger the disease processes. A third requirement for the concept to work is the capability to obtain samples from control unperturbed disease tissues and from the perturbed disease tissues and obtain gene and protein information, such that the changes in genes and proteins that occur as a result

of the perturbation can be discriminated. The final requirement is for a method that can easily test identified genes and proteins for the ones that were actually driving the response to the perturbation (i.e., the disease efficacy process), as these are the drug targets. The combination of the discovery process, which is shown in Figure 3, with the final validation step is shown in Figure 4. Each of these four requirements is considered separately.

Gene Selection

Basically, the concept is to select genes such that their perturbation in the disease generates the desired disease dynamics. For cancer, the desired change is an altered inhibited growth rate. A number of genes can be inhibited or expressed in the tumor or elsewhere in the animal that alter animal model tumor growth rate. The result is a system that provides tumors with stimulated and inhibited growth rates, as well as progressing tumors from negative control treatments. Analysis of the two types of tumors, control and inhibited, allows characterization of the gene and protein differential between control tumor and inhibited tumor. One element of the system is the biological amplification of gatekeepers upon the induction of stimulation or inhibition of tumor growth. This amplification results in a significant enrichment of the data set generated with gatekeeper genes and proteins. Note that equally valuable information results from stimulation as from inhibition of tumor growth rate. It seems that either accelerated or inhibited growth rate dynamics should provide the key genes and proteins associated with tumor growth dynamics. In fact, the most interesting gene selection is where two different gene perturbations are used separately, such that tumor growth rate is accelerated in one case and inhibited in the other. This way, tumors can be obtained that have responded in both directions of the dynamic process of interest.

Unknown genes or suspected key genes can be selected for altered expression in the disease rather than known effector genes. In this case, and in particular for suspected key genes, their perturbation in the active disease is essentially a test of whether they have a gatekeeper role in the pathology and whether that role is appropriate for inhibitors, antagonists, or for activators, agonists. More importantly, combination with pathway analysis on the treated disease tissues can enable identification of associated genes and proteins downstream or even upstream of the suspect, which are new candidates for gatekeeper function in the disease process.

Perturbations other than altered gene expression can be used to induce disease dynamics. Many clinically relevant preclinical disease models have been used to obtain efficacy evidence for clinical drug development, many of which are approved for human use. Thus, these can be used to obtain efficacy disease dynamics. However, virtually all drugs in use today have been identified by empirical methods rather than engineered inhibitors of a specific drug target. Their pathway effects are pleotropic rather than a very specific effect of a single gene. The combination of a gene perturbation with drug treatment can be used to generate biological information on the pathway(s) in which the drug is acting, both to induce efficacy and to induce toxicities. This appears to represent a powerful approach to use existing drugs in combination with known pathway effector genes, thereby generating biological data sets rich in dynamic pathway information.

Nucleic Acid Delivery: Gene Expression or Inhibition

The concept depends on nucleic acid delivery that is sufficiently effective to significantly alter gene expression and thereby induce macroscopic pharmacological changes in tumor growth. Effective nucleic gene delivery in the animal tumor model, which is activity in a significant number of cells, is crucial, but an equally important requirement is delivery without background pharmacological activity from the delivery method itself. Lipid- and polymer-based delivery systems have been developed to provide nucleic acid activity at many tissues, including tumors, and in some cases with relatively low background for solid tumors. Viral vector systems also have been developed to provide nucleic acid activity at tumors, but tend to have stronger background from simultaneous biological activities of the viral protein. Also, viral vectors are limited to natural nucleic acids, so that chemically modified nucleic acid analogues require physical and chemical delivery methods.

Extensive work with oligonucleotide gene inhibitors (antisense, ribozymes, etc.) and gene therapy has demonstrated a variety of nucleic acid delivery methods giving sufficiently effective activity to produce animal tumor model efficacy. In some cases, the nucleic acids have been administered locally, frequently directly into the tumor, confining

the effect to the local tissues. In other cases, they have been systemically administered, although typically with a less selective effect at the tumor. While these preclinical successes of antisense and gene therapy have motivated clinical trials, the delivery methods have yet to provide any clinical success. Consequently, continued preclinical efforts are directed toward more effective delivery, newer targets, and newer forms of nucleic acids (e.g., recent work with DNAzyme inhibitors of vascular endothelial growth factor [VEGF] receptors) (39).

Recent work recognized the usefulness of systemic gene expression in preclinical tumor models for the determination of gene function rather than as a therapeutic approach (38). In this work, nucleic acid delivery to the liver was able to achieve strong expression of proteins and systemic levels. The systemic activity of the expressed proteins was sufficient to verify gene and protein function. A major advantage of this method is the separation of nucleic acid delivery and gene expression (i.e., liver) from the subcutaneous tumor tissue. Thus, any pharmacological effects of the nucleic acid delivery and expression in liver can be avoided in genomic and proteomic readouts from the tumor, which is perturbed by systemically acting proteins. This specific success of gene expression for gene function validation in preclinical tumor models provides assurance of sufficient effectiveness to be used for the concept but is limited to systemically acting factors.

For gene expression or inhibition confined to the tumor tissue, much of the recent gene therapy efforts are focused on local administration of tumor selective viral replication (36,37), which achieves a macroscopic effect on tumor growth but by viral replication and tumor cell killing rather than specific gene expression. The development of delivery methods that can achieve tumor-targeted nucleic acid delivery and activity directly within the tumor has been limiting. A recent report using ligand receptor targeting specific for tumor neovasculature suggested that the method could be developed for clinical therapeutics (40), but nonetheless provides very strong evidence for preclinical tumor model nucleic acid delivery sufficient for verifying gene function by overexpressing a negative dominant signal factor. The avoidance of effects from the nucleic acid delivery acting at the tumor rather than the gene activity still needs to be minimized. Consequently, the synthetic delivery methods that tend to have lower levels of effects are preferred relative to viral vectors with many protein components with a tendency for pharmacological effects of the their own. Overall, the successes obtaining macroscopic effects on preclinical model tumor growth rates by delivery of oligonucleotide inhibitors and expression of genes, which is work undertaken largely with the goal of developing potential clinical therapeutics, can be adapted for gene perturbation of tumor growth rate as a biological system for genomics and proteomics. Continued improvements in efficiency and effectiveness of nucleic acid delivery in preclinical animal models will improve their application for gene function studies. In conclusion, the effectiveness of nucleic acid delivery in preclinical tumor models appears to be adequate, but represents the aspect that can benefit the most from continued development.

Perturbed Tissue Pathway Analysis: Genomics and Proteomics

Virtually any and all of the large and growing number of methods available to analyze gene and protein activity in biological samples can be used with perturbed tumor tissues. The refinement of these methods has improved considerably, both in throughput and in sensitivity and resolution. Any method that has been developed for tissue sections from human clinical samples or animal disease models can and should be considered for obtaining genomic and proteomic information from animal tumor models perturbed by gene expression or inhibition. In fact, a combination of analyses should be used to obtain data sets with the greatest content. Moreover, in addition to genomic and proteomic analyses, pharmacological and pathological characterizations of the tumor can and should be obtained for inclusion in the data set. Based on a combination of genomic or proteomic methods with their bioinformatic analyses with traditional biological analysis, the genes and proteins significantly up- or down-regulated in the defined pathway can be carefully selected and further analyzed. This presents a challenge to the design of the data set in that such pharmacological and pathological characterizations are often in a different form from the genomic and proteomic array data. Often these characterizations are recorded as "annotation". Improved data set structures will facilitate the inclusion and use of widely varying forms of data. Advances in combining and sorting such disparate forms of data will be tremendously helpful with the use of preclinical animal models.

The use of both stimulating and inhibiting perturbations, possibly through the use of gene

perturbations with opposing effects, makes it possible to identify the genes and proteins that correlate with both dynamic changes in the disease. In cancer, this means enhanced and inhibited tumor growth rate. Thus, use of genomic and proteomic methods that can identify genes and proteins whose movement matches that of both changes in tumor growth rate appear to be most useful. Criteria for identification of the genes and proteins depend on the method employed. Note that as new measures of biological systems become available, such as carbohydrate modifications of proteins, these measures can and should be utilized.

Finally, the selection of broad or narrow pathway analysis method needs to be considered. Narrow pathway analysis methods can be selected for specific pathological and pathway information. On the other hand, genome-wide pathway analysis methods can be used, since the differential between control tumor and perturbed tumor is expected to yield only small numbers of gene and protein differences. Therefore, the sensitivity of the methods needs to be high. This is likely true for another more fundamental reason, namely the levels of gatekeepers tends to be low relative to the overall population. The gatekeepers generally initiate or block cascades, so their levels tend to be low in comparison to the members of the cascade, especially the final stage of the cascade. A combination of sensitive genome- or proteome-wide pathway analysis methods with gene perturbation provides disease control-focused discovery.

Disease-Control Validation

The data sets produced by application of pathway analysis to compare perturbed tumors with control tumors contain the genes and proteins that associate and correlate with the perturbation but still require sorting (i.e., validation) for the ability to control the perturbation. A comparison of expression data may reveal, for example, that a particular gene's expression decreases with inhibited tumor growth rates, suggesting it's inhibition may inhibit tumor growth. Inhibiting that gene selectively will illustrate if it is in fact a controlling element of tumor growth rate, a drug target, and whether it is stronger or weaker than the gene used to perturb the tumor in the first place. These newly identified candidate targets may have only speculative functions or be unknown. Complete validation depends on two issues: (*i*) whether they have gatekeeper properties that control the disease; and (*ii*) whether they are selective for the disease pathology or are required in healthy biological processes. Sorting candidates first for disease control should provide a small set of candidates for testing for toxicity and is much more efficient than in reverse.

As in identification, validation can also be accelerated by the use of preclinical animal disease models. In one straightforward approach, the nucleic acid delivery used to alter expression of the perturbing gene in the preclinical tumor model can be used to perturb expression of the candidate target. After a period of time, the animal tumor growth rate can be analyzed to reveal whether the candidate gene significantly controls the disease. All the same requirements for the nucleic acid delivery used for the discovery discussed above apply here. A comparison with the appropriate positive and negative control genes can reveal whether the candidate gene is a disease-controlling gene and the relative strength of its controlling capabilities. This animal tumor model validation has been described using liver expression of systemically active cytokines (38). The continuing advances in nucleic acid delivery, along with improved sequence-specific oligonucleotide inhibitors, such as RNAi, are addressing this key technological hurdle for direct animal model validation of targets for efficacy.

Other aspects beyond disease-controlling capabilities are important for validating a drug target. Consequently, further validation for each newly identified and validated target is required (e.g., different tumor models, biochemistry and cell biology, and histology) to justify the next step of drug discovery process. For example, it is important to determine the cell pathways of each target and their expression levels in healthy tissues versus disease tissues. The many high-throughput screen data sets already being assembled contain much of the critical information needed for these questions. For example, expression profiling of thousands of human tissue samples characterized for pathological and normal states can be used to select those disease-controlling targets with low prevalence in healthy tissues so as to minimize the toxicity of drugs selectively inhibiting that target. While these aspects can take considerable efforts to assess, the efforts can be focused on those genes and proteins with known disease-controlling capabilities (i.e., a narrow bandwidth) and as a result be much more valuable.

CONCLUSION

The use of preclinical animal models of disease, and in particular cancer models, promises to revolutionize drug target discovery. They can be used to accelerate efforts to identify the genes and proteins best suited as drug targets by focusing studies on the salient aspect, such as disease control. Their importance comes about from their unique capability to provide information on the critical issue, namely control of established disease processes (i.e., efficacy). Advances in nucleic acid delivery into tumors and other tissues of these animal models are providing the enabling capability for this revolution. First, they permit efficacy processes to be induced in established tumor models via perturbation of specific genes and proteins, resulting in tumor growth rate changes that can be analyzed with genomic and proteomic methods to reveal the genes and proteins correlating with efficacy. Second, they permit direct sorting of the identified genes and proteins for their own ability to control the disease process and yield efficacy. These targets, validated for efficacy, can then be studied with the standard genomic and proteomic methods in a narrow focused manner to complete their validation as cancer drug targets.

ACKNOWLEDGMENTS

The authors appreciate many helpful discussions with Drs. Robert W. Malone, Victor Haodong Li, Dexi Liu, Anton Wellstein, and Wei-Wu He.

REFERENCES

1. **Syzmkowski, D.E.** 2001. Too many targets, not enought target validation. Drug Discov. Today *6*:398-399.
2. **Knowles, J. and G. Gromo.** 2003. A guide to drug discovery: target selection in drug discovery. Nat. Rev. Drug Discov. *2*:63-69.
3. **Zhang, M.Q.** 1999. Large-scale gene expression data analysis: a new challenge to computational biologists. Genome Res. *9*:681-688.
4. **Ohlstein, E.H., R.R. Ruffolo, Jr., and J.D. Elliott.** 2000. Drug discovery in the next millennium. Ann. Rev. Pharmacol. Toxicol. *40*: 177-191.
5. **Loewe, L.** 2002. Global computing for bioinformatics. Brief Bioinform. *3*:377-388.
6. **Orr, M.S. and U. Scherf.** 2002. Large-scale gene expression analysis in molecular target discovery. Leukemia *16*:473-477.
7. **Ostermeier, G.C., D.J. Dix, and S.A. Krawetz.** 2002. A bioinformatic strategy to rapidly characterize cDNA libraries. Bioinformatics *18*:949-952.
8. **Jenkins, J.L., R.Y. Kao, and R. Shapiro.** 2003. Virtual screening to enrich hit lists from high-throughput screening: a case study on small-molecule inhibitors of angiogenin. Proteins *50*:81-93.
9. **Cho, C.H. and M.E. Nuttall.** 2002. Emerging techniques for the discovery and validation of therapeutic targets for skeletal diseases. Expert Opin. Ther. Targets *6*:679-689.
10. **Lee, D., S.A. Long, J.L. Adams, G. Chan, K.S. Vaidya, T.A. Francis, K. Kikly, J.D. Winkler, et al.** 2000. Potent and selective nonpeptide inhibitors of caspases 3 and 7 inhibit apoptosis and maintain cell functionality. J. Biol. Chem. *270*:16007-16014.
11. **Nuttall, M.E.** 2001. Drug discovery and target validation. Cells Tissues Organs *169*:265-271.
12. **Kitano, H.** 2002. Computational systems biology. Nature *420*:206-210.
13. **Aravind, L. and L.M. Iyer.** 2002. Intraproteomic networks: new forays into predicting interaction partners. Genome Res *12*:1156-1158.
14. **Calvo, A., N. Xiao, J. Kang, C.J. Best, I. Leiva, M.R. Emmert-Buck, C. Jorcyk, and J.E. Green.** 2002. Alterations in gene expression profiles during prostate cancer progression: functional correlations in tumorigenicity and down-regulation of selenoprotein-P in mouse and human tumors. Cancer Res. *62*:5325-5335.
15. **Stumm, G., A. Russ, and M. Nehls.** 2002. Deductive genomics: a functional approach to identify innovative drug targets in the post-genome era. Am. J. Pharmacogenomics *2*:263-271.
16. **Kim, R., A. Trubetskoy, T. Suzuki, N.A. Jenkins, N.G. Copeland, and J. Lenz.** 2003. Genome-based identification of cancer genes by proviral tagging in mouse retrovirus-induced T-cell lymphomas. J. Virol. *77*:2056-2062.
17. **Kurella, M., L.L. Hsiao, T. Yoshida, J.D. Randall, G. Chow, S.S. Sarang, R.V. Jensen, and S.R. Gullans.** 2001. DNA microarray analysis of complex biological processes. J. Am. Soc. Nephrol. *12*:1072-1078.
18. **Shoemaker, D.D. and P.S. Linsley.** 2002. Recent developments in DNA microarrays. Curr. Opin. Microbiol. *5*:334-337.
19. **Gerritsen, M.E., F.V. Peale, Jr., and T. Wu.** 2002. Gene expression profiling in silico: relative expression of candidate angiogenesis associated genes in renal cell carcinomas. Exp. Nephrol. *10*:114-119.
20. **Hanash, S.M., J. Madoz-Gurpide, and D.E. Misek.** 2002. Identification of novel targets for cancer therapy using expression proteomics. Leukemia *16*:478-485.
21. **Chapman, K.** 2002. The ProteinChip Biomarker System from Ciphergen Biosystems: a novel proteomics platform for rapid biomarker discovery and validation. Biochem. Soc. Trans. *30*:82-87.
22. **Emmert-Buck, M.R., R.F. Bonner, P.D. Smith, R.F. Chuaqui, Z. Zhuang, S.R. Goldstein, R.A. Weiss, and L.A. Liotta.** 1996. Laser capture microdissection. Science *274*:998-1001.
23. **Bichsel, V.E., L.A. Liotta, and E.F. Petricoin, 3rd.** 2001. Cancer proteomics: from biomarker discovery to signal pathway profiling. Cancer J. *7*:69-78.
24. **Elkahloun, A.G., J. Gaudet, G.S. Robinson, and D.C. Sgroi.** 2002. In situ gene expression analysis of cancer using laser capture microdissection, microarrays and real time quantitative PCR. Cancer Biol. Ther. *1*:354-358.
25. **Eltoum, I.A., G.P. Siegal, and A.R. Frost.** 2002. Microdissection of histologic sections: past, present, and future. Adv. Anat. Pathol. *9*:316-322.
26. **Brady, G.** 2000. Expression profiling of single mammalian cells—small is beautiful. Yeast *17*:211-217.
27. **Seth, D., M.D. Gorrell, P.H. McGuinness, M.A. Leo, C.S. Lieber, G.W. McCaughan, and P.S. Haber.** 2003. SMART amplification maintains representation of relative gene expression: quantitative validation by real time PCR and application to studies of alcoholic liver disease in primates. J. Biochem. Biophys. Methods *55*:53-66.
28. **Takahashi, J.S., L.H. Pinto, and M.H. Vitaterna.** 1994. Forward and reverse genetic approaches to behavior in the mouse. Science *264*: 1724-1733.
29. **Kay, S.** Using mouse forward genetics to create novel target space. Genomics on Target, Functional Genomics, Nov 18-21, 2002. Cambridge Healthtech Institute, Sheraton Boston Hotel, Boston, MA.
30. **Walke, D.W., C. Han, J. Shaw, E. Wann, B. Zasmbrowicz, and A. Sands.** 2001. In vivo drug target discovery: identifying the best targets from the genome. Curr. Opin. Biotech. *12*:626-631.

31. **Harris, S.** 2001. Transgenic knockouts as part of high-throughput, evidence-based target selection and validation strategies. Drug Discov. Today *6*:398-399.
32. **Bockamp, E., M. Maringer, C. Spangenberg, S. Fees, S. Fraser, L. Eshkind, F. Oesch, and B. Zabel.** 2002. Of mice and models: improved animal models for biomedical research. Physiol. Genomics *11*:115-132.
33. **Dias, N. and C.A. Stein.** 2002. Antisense oligonucleotides: basic concepts and mechanisms. Mol. Cancer Ther. *1*:347-355.
34. **Stein, C.A.** 2001. The experimental use of antisense oligonucleotides: a guide for the perplexed. J. Clin. Invest. *108*:641-644.
35. **Xia, H., Q. Mao, H.L. Paulson, and B.L. Davidson.** 2002. siRNA-mediated gene silencing in vitro and in vivo. Nat. Biotechnol. *20*: 1006-1010.
36. **XI, S. and J.R. Grandis.** 2003. Gene therapy for the treatment of oral squamous cell carcinoma. J. Dent. Res. *82*:11-16.
37. **LI, S., J. Simons, N. Detorie, B. O'Rourke, U. Hamper, and T.L. DeWeese.** 2003. Dosimetric and technical considerations for interstitial adenoviral gene therapy as applied to prostate cancer. Int. J. Radiat. Oncol. Biol. Phys. *55*:204-214.
38. **Liu, D. and J.E. Knapp.** 2001. Hydrodynamics-based gene delivery. Curr. Opin. Mol. Ther. *3*:192-197.
39. **Zhang, L., W.J. Gasper, S.A. Stass, O.B. Ioffe, M.A. Davis, and A.J. Mixson.** 2002. Angiogenic inhibition mediated by a DNAzyme that targets vascular endothelial growth factor receptor 2. Cancer Res. *62*: 5463-5469.
40. **Hood, J.D., M. Bednarski, R. Frausto, S. Guccione, R.A. Reisfeld, R. Xiang, and D.A. Cheresh.** 2002. Tumor regression by targeted gene delivery to the neovasculature. Science *296*:2404-2407.

Address correspondence to:

Dr. Martin Woodle
Intradigm Corporation
12115 K Parklawn Drive
Rockville, MD 20852, USA
e-mail: mwoodle@intradigm.com

OVERVIEW TALKS

THE IMPORTANCE OF BEING ABLE TO PREDICT HUMAN SENSITIVITY

Joseph E. Tomaszewski, Ph.D.
National Cancer Institute

In oncology, where many patients are treated with toxic agents, the ability to predict human sensitivity is of paramount importance. With terminal cancer patients as subjects in phase I clinical trials, it is important to have a good paradigm to predict toxicity. This presentation examined time and financial considerations, toxicity considerations, predictability of animal studies, predictability of *in vitro* studies, and a wish list for future research.

The road from drug discovery to approval is long and arduous and can take 15–16 years from discovery to marketing. Of 5,000–10,000 or more compounds that are initially screened, 2,500 enter preclinical testing, five enter clinical testing, and one is ultimately approved by the Food and Drug Administration (FDA). Typically, it costs $800 million to bring a new drug to market. Including anti-infectives, 46 percent of drug development is terminated because of lack of efficacy, 17 percent because of animal toxicity, 16 percent because of adverse reactions in humans, and seven percent for either pharmakinetic, commercial, or miscellaneous reasons.

Serious liver injury is the primary reason for drug removals or restrictions. A recent examination of drugs withdrawn from the market for unexpected adverse events in a 12–month period raised questions about the FDA's drug review process. Some (for example, fenfluramine) had been on the market for years with millions of patient exposures before being withdrawn. Some had relatively few patients exposed to the product during testing, and it took the many post-market cases to detect problems.

Toxicities, such as seen with troglitazone, an oral hypoglycemic agent, and Baycol, a statin, probably could not have been predicted through traditional toxicity studies, although drug/drug interaction studies might have predicted some. Animal tests increase the predictability of human response and show bone marrow and gastrointestinal toxicities, but other toxicities are not that well predicted. The mouse is the poorest predictor of human toxicity; dog and nonhuman primates yield better results. For example, dogs were better than mice and mice better than rats in predicting maximum tolerated dose. However, dogs and nonprimate humans are difficult to get and expensive.

Looking at the value of *in vitro* data, in studies by NCI and others, bone marrow assay results found that mouse data alone accurately predicted human maximum dose for 38 of 48 drugs, and mouse and dog data accurately predicted for 43 of 48 drugs. There is disparity between species in bone marrow data with much more sensitivity seen in humans, but researchers have determined that in general *in vitro* bone marrow assays are highly predictive of human sensitivity.

For future research, scientists should aim not for assays that merely rank new chemical entities or assays with concentrations that are not physiologically relevant, but rather for quantitative *in vitro* assays that are predictive of human sensitivity in relation to animal sensitivity, targeting the brain, GI tract, heart, kidney, liver, and lung. Ideally, *in vivo* and *in vitro* animal and human data

from tissue banks will be available to generate accurate predictions for starting doses, maximum tolerated dose, dose limiting toxicity, and pharmacokinetics.

TOXICOLOGY LIABILITIES AND THEIR EFFECT ON SUCCESS RATES IN THE PHARMACEUTICAL INDUSTRY: HOW CAN WE SHARPEN THE TOOLS THAT WE HAVE OR FORCE NEW ONES?
Richard Robertson, Ph.D.
Bristol Meyer Squibb Company

The industry dilemma is that as research and development (R&D) investment has increased, the number of approvals has declined. The problem is real—while the mid–1980s to the mid–1990s are referred to as the golden age of pharmaceuticals, the situation is now a perilous morass, with companies consuming each other and a resultant decreasing number of companies actually engaged in pharmaceutical research.

While the success rate from submission to market had an upturn in the past ten years, with an approximate 90 percent success rate, the success rates from first pivotal dose have dropped from 70–50 percent. Large declines are also seen in first patient dose to market (30–12 percent) and first human dose to market, which is down to ten percent, a rate that is not sustainable.

Large differences are seen in success rates in the various therapeutic classes. There are much higher rates in the anti-infectives because mechanisms are understood, with surrogate markers from *in vitro* data. But cardiovascular and anti-cancer drugs have much lower success rates, and success for nervous system drugs is worst of all, with an abysmal one percent of nervous system drugs that enter human trials making it to the market. There are many reasons for this, including lack of good animal drugs and the difficulty in designing good clinical protocols. This is an area of huge unmet medical need.

The declining success rates result from a confluence of factors: the demand for larger safety bases; the rapidly escalating cost of clinical development programs with downward pressures on drug pricing; the vastly expanded knowledge of the human genome, which allows identification of targets to outstrip the ability to understand and develop them; and the inability to transfer surrogate markers to clinical use and to find the right target and understand its biology.

Several measures can address the problems. Scientists need to find the right target and understand its biology. Then the task is to find the right molecule—its pharmacology, ADME, toxicologic liabilities, and pharmaceutical properties. It is necessary to develop a system to fail early and often to allow as many shots as possible at the goal—predictive, high throughput, integrated, interactive screening across multiple systems pharmacology, ADME, toxicology, and pharmaceutics. The tools to do this are not yet in place, but it is where the industry needs to go.

Specifically, this means improved target organ-specific knockout technologies in mice and rats, precise examination of the role of biotransformation in target organ drug toxicity, and hypothesis testing of the role of specific gene expression pathways to organ toxicity. It means major improvements in the use of primary cell cultures stable cell lines for predictive toxicology assays, with the ability to maintain cellular phenotypes *in vitro* to allow direct comparison of possible target organ toxicities and mechanisms between laboratory animals and humans. Biology systems development is needed to develop pathway tools for proteomics—tools linking mRNA, protein, and intermediary metabolic changes with chemical scaffolds would allow toxicology hypothesis development and testing. Improved proteomics platforms are also